**HW5 – Theory + SVM**

1. PAC Learning and VC dimension (30 pts)

Let . Let   
, for ,

the set of all origin-centered rings.

* 1. (8 pts) What is the ? Prove your answer.
  2. (14 pts) Describe a polynomial sample complexity algorithm that learns using . State the time complexity and the sample complexity of your suggested algorithm. Prove all your steps.

In class we saw a bound on the sample complexity when is finite.  
When is infinite, we have a different bound:

* 1. (8 pts) You want to get with 95% confidence a hypothesis with at most 5% error. Calculate the sample complexity with the bound that you found in b and the above bound for infinite . In which one did you get a smaller ? Explain.

**Handwritten answer in next page**

1. VC dimension (20 pts)

Let and .

Define “x-node decision tree” for any to be a full binary decision tree with x nodes (including the leaves).   
Let be the hypothesis space of all “x-node decision tree” with .

* 1. (5 pts) What is the ? Prove your answer.
  2. (15 pts) What is the ? Prove your answer.

**Handwritten answer in next page**

1. Kernels and mapping functions (25 pts)
   1. (20 pts) Let be a function over (i.e., ).

Find for which is a kernel. (It may help to first expand the above term on the right-hand side).

**Answer:**

We’ll use the following binomial theorem formulas to expand Kernel function:

Each term in this expansion represents one element in the feature mapping , So K is a kernel of:

* 1. (2 pts) What did we call the function in class if we remove all coefficients?

**Answer:**

If we remove all coefficients we will have:

**We called this a full rational variety of order 3**

This is in an input space of dimension 2 is described by all 3-th degree monomials of the input variables in x.

The size of this group should be while here and so we have items here.

* 1. (3 pts) How many multiplication operations do we save by using versus ?  
     **Answer:**

Using K, we will have 2 multiplications for the dot product over

Using , we will have 10 multiplications for the dot product of 10 elements in the feature space.

Therefore, we will save multiplication operations by using the Kernel function.

1. Lagrange multipliers (15 pts)

Let . Find the minimum and the maximum points for under the constraint .

**Answer:**

To find the minimum and maximum points using Lagrange multipliers, we define the Lagrangian function:

T**he critical points are:**

**So, our finals results are:**

1. See notebook exercise (10 pts)

**Answer in the attached Jupyter notebook**